

#10

1652

RAW SEQUENCE LISTING

DATE: 09/24/2001

PATENT APPLICATION: US/09/495,823

TIME: 12:05:32

Input Set : A:\5800-79.txt

Output Set: N:\CRF3\09242001\I495823.raw

4 <110> APPLICANT: Glucksman, Maria Alexandra
 5 Williamson, Mark
 8 <120> TITLE OF INVENTION: 22438, 23553, 25278, and 26212 Novel
 9 Human Sulfatases
 11 <130> FILE REFERENCE: 35800/191890(5800-79
 13 <140> CURRENT APPLICATION NUMBER: 09/495,823
 14 <141> CURRENT FILING DATE: 2000-01-31
 16 <160> NUMBER OF SEQ ID NOS: 8
 18 <170> SOFTWARE: FastSEQ for Windows Version 4.0
 20 <210> SEQ ID NO: 1
 21 <211> LENGTH: 525
 22 <212> TYPE: PRT
 23 <213> ORGANISM: homo sapiens
 25 <400> SEQUENCE: 1
 26 Met Gly Trp Leu Phe Leu Lys Val Leu Leu Ala Gly Val Ser Phe Ser
 27 1 5 10 15
 28 Gly Phe Leu Tyr Pro Leu Val Asp Phe Cys Ile Ser Gly Lys Thr Arg
 29 20 25 30
 30 Gly Gln Lys Pro Asn Phe Val Ile Ile Leu Ala Asp Asp Met Gly Trp
 31 35 40 45
 32 Gly Asp Leu Gly Ala Asn Trp Ala Glu Thr Lys Asp Thr Ala Asn Leu
 33 50 55 60
 34 Asp Lys Met Ala Ser Glu Gly Met Arg Phe Val Asp Phe His Ala Ala
 35 65 70 75 80
 36 Ala Ser Thr Cys Ser Pro Ser Arg Ala Ser Leu Leu Thr Gly Arg Leu
 37 85 90 95
 38 Gly Leu Arg Asn Gly Val Thr Arg Asn Phe Ala Val Thr Ser Val Gly
 39 100 105 110
 40 Gly Leu Pro Leu Asn Glu Thr Thr Leu Ala Glu Val Leu Gln Gln Ala
 41 115 120 125
 42 Gly Tyr Val Thr Gly Ile Ile Gly Lys Trp His Leu Gly His His Gly
 43 130 135 140
 44 Ser Tyr His Pro Asn Phe Arg Gly Phe Asp Tyr Tyr Phe Gly Ile Pro
 45 145 150 155 160
 46 Tyr Ser His Asp Met Gly Cys Thr Asp Thr Pro Gly Tyr Asn His Pro
 47 165 170 175
 48 Pro Cys Pro Ala Cys Pro Gln Gly Asp Gly Pro Ser Arg Asn Leu Gln
 49 180 185 190
 50 Arg Asp Cys Tyr Thr Asp Val Ala Leu Pro Leu Tyr Glu Asn Leu Asn
 51 195 200 205
 52 Ile Val Glu Gln Pro Val Asn Leu Ser Ser Leu Ala Gln Lys Tyr Ala
 53 210 215 220
 54 Glu Lys Ala Thr Gln Phe Ile Gln Arg Ala Ser Thr Ser Gly Arg Pro
 55 225 230 235 240
 56 Phe Leu Leu Tyr Val Ala Leu Ala His Met His Val Pro Leu Pro Val
 57 245 250 255
 58 Thr Gln Leu Pro Ala Ala Pro Arg Gly Arg Ser Leu Tyr Gly Ala Gly

ENTERED

See page 5

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/495,823

DATE: 09/24/2001

TIME: 12:05:32

Input Set : A:\5800-79.txt

Output Set: N:\CRF3\09242001\I495823.raw

```

59          260          265          270
60 Leu Trp Glu Met Asp Ser Leu Val Gly Gln Ile Lys Asp Lys Val Asp
61          275          280          285
62 His Thr Val Lys Glu Asn Thr Phe Leu Trp Phe Thr Gly Asp Asn Gly
63          290          295          300
64 Pro Trp Ala Gln Lys Cys Glu Leu Ala Gly Ser Val Gly Pro Phe Thr
65 305          310          315          320
66 Gly Phe Trp Gln Thr Arg Gln Gly Gly Ser Pro Ala Lys Gln Thr Thr
67          325          330          335
68 Trp Glu Gly Gly His Arg Val Pro Ala Leu Ala Tyr Trp Pro Gly Arg
69          340          345          350
70 Val Pro Val Asn Val Thr Ser Thr Ala Leu Leu Ser Val Leu Asp Ile
71          355          360          365
72 Phe Pro Thr Val Val Ala Leu Ala Gln Ala Ser Leu Pro Gln Gly Arg
73          370          375          380
74 Arg Phe Asp Gly Val Asp Val Ser Glu Val Leu Phe Gly Arg Ser Gln
75 385          390          395          400
76 Pro Gly His Arg Val Leu Phe His Pro Asn Ser Gly Ala Ala Gly Glu
77          405          410          415
78 Phe Gly Ala Leu Gln Thr Val Arg Leu Glu Arg Tyr Lys Ala Phe Tyr
79          420          425          430
80 Ile Thr Gly Gly Ala Arg Ala Cys Asp Gly Ser Thr Gly Pro Glu Leu
81          435          440          445
82 Gln His Lys Phe Pro Leu Ile Phe Asn Leu Glu Asp Asp Thr Ala Glu
83          450          455          460
84 Ala Val Pro Leu Glu Arg Gly Gly Ala Glu Tyr Gln Ala Val Leu Pro
85 465          470          475          480
86 Glu Val Arg Lys Val Leu Ala Asp Val Leu Gln Asp Ile Ala Asn Asp
87          485          490          495
88 Asn Ile Ser Ser Ala Asp Tyr Thr Gln Asp Pro Ser Val Thr Pro Cys
89          500          505          510
90 Cys Asn Pro Tyr Gln Ile Ala Cys Arg Cys Gln Ala Ala
91          515          520          525
94 <210> SEQ ID NO: 2
95 <211> LENGTH: 2175
96 <212> TYPE: DNA
97 <213> ORGANISM: homo sapiens
99 <220> FEATURE:
100 <221> NAME/KEY: CDS
101 <222> LOCATION: (248)...(1825)
103 <400> SEQUENCE: 2
104 cacgogtccg caaatttccct gattcttttg aattaggatt ccagatgggg gcctcatttc 60
105 tacagccccc aacatttcta tagccgttat cactgccatc accactgccca ccagcatctt 120
106 cttgcagatt ccaccctgc tccccagaga ctctctgctt tgaaagtgag cagaaaggaa 180
107 gctctcagaa aaatctctag tgggtgctgc cgctcgtcca gacaatcgga atcctgcctt 240
108 caccacc atg ggc tgg ctt ttt cta aag gtt ttg ttg gcg gga gtg agt 289
109          Met Gly Trp Leu Phe Leu Lys Val Leu Leu Ala Gly Val Ser
110          1          5          10
112 ttc tca gga ttt ctt tat cct ctt gtg gat ttt tgc atc agt ggg aaa 337

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/495,823

DATE: 09/24/2001

TIME: 12:05:32

Input Set : A:\5800-79.txt

Output Set: N:\CRF3\09242001\I495823.raw

113	Phe	Ser	Gly	Phe	Leu	Tyr	Pro	Leu	Val	Asp	Phe	Cys	Ile	Ser	Gly	Lys	
114	15					20					25					30	
116	aca	aga	gga	cag	aag	cca	aac	ttt	gtg	att	att	ttg	gcc	gat	gac	atg	385
117	Thr	Arg	Gly	Gln	Lys	Pro	Asn	Phe	Val	Ile	Ile	Leu	Ala	Asp	Asp	Met	
118					35					40					45		
120	ggg	tgg	ggt	gac	ctg	gga	gca	aac	tgg	gca	gaa	aca	aag	gac	act	gcc	433
121	Gly	Trp	Gly	Asp	Leu	Gly	Ala	Asn	Trp	Ala	Glu	Thr	Lys	Asp	Thr	Ala	
122				50					55					60			
124	aac	ctt	gat	aag	atg	gct	tcg	gag	gga	atg	agg	ttt	gtg	gat	ttc	cat	481
125	Asn	Leu	Asp	Lys	Met	Ala	Ser	Glu	Gly	Met	Arg	Phe	Val	Asp	Phe	His	
126			65					70					75				
128	gca	gct	gcc	tcc	acc	tgc	tca	ccc	tcc	cgg	gct	tcc	ttg	ctc	acc	ggc	529
129	Ala	Ala	Ala	Ser	Thr	Cys	Ser	Pro	Ser	Arg	Ala	Ser	Leu	Leu	Thr	Gly	
130		80						85				90					
132	cgg	ctt	ggc	ctt	cgc	aat	gga	gtc	aca	cgc	aac	ttt	gca	gtc	act	tct	577
133	Arg	Leu	Gly	Leu	Arg	Asn	Gly	Val	Thr	Arg	Asn	Phe	Ala	Val	Thr	Ser	
134	95					100				105						110	
136	gtg	gga	ggc	ctt	ccg	ctc	aac	gag	acc	acc	ttg	gca	gag	gtg	ctg	cag	625
137	Val	Gly	Gly	Leu	Pro	Leu	Asn	Glu	Thr	Thr	Leu	Ala	Glu	Val	Leu	Gln	
138					115					120					125		
140	cag	gcg	ggt	tac	gtc	act	ggg	ata	ata	ggc	aaa	tgg	cat	ctt	gga	cac	673
141	Gln	Ala	Gly	Tyr	Val	Thr	Gly	Ile	Ile	Gly	Lys	Trp	His	Leu	Gly	His	
142				130						135				140			
144	cac	ggc	tct	tat	cac	ccc	aac	ttc	cgt	ggt	ttt	gat	tac	tac	ttt	gga	721
145	His	Gly	Ser	Tyr	His	Pro	Asn	Phe	Arg	Gly	Phe	Asp	Tyr	Tyr	Phe	Gly	
146			145					150				155					
148	atc	cca	tat	agc	cat	gat	atg	ggc	tgt	act	gat	act	cca	ggc	tac	aac	769
149	Ile	Pro	Tyr	Ser	His	Asp	Met	Gly	Cys	Thr	Asp	Thr	Pro	Gly	Tyr	Asn	
150		160					165					170					
152	cac	cct	cct	tgt	cca	gcg	tgt	cca	cag	ggt	gat	gga	cca	tca	agg	aac	817
153	His	Pro	Pro	Cys	Pro	Ala	Cys	Pro	Gln	Gly	Asp	Gly	Pro	Ser	Arg	Asn	
154	175					180					185					190	
156	ctt	caa	aga	gac	tgt	tac	act	gac	gtg	gcc	ctc	cct	ctt	tat	gaa	aac	865
157	Leu	Gln	Arg	Asp	Cys	Tyr	Thr	Asp	Val	Ala	Leu	Pro	Leu	Tyr	Glu	Asn	
158				195						200				205			
160	ctc	aac	att	gtg	gag	cag	ccg	gtg	aac	ttg	agc	agc	ctt	gcc	cag	aag	913
161	Leu	Asn	Ile	Val	Glu	Gln	Pro	Val	Asn	Leu	Ser	Ser	Leu	Ala	Gln	Lys	
162				210					215					220			
164	tat	gct	gag	aaa	gca	acc	cag	ttc	atc	cag	cgt	gca	agc	acc	agc	ggg	961
165	Tyr	Ala	Glu	Lys	Ala	Thr	Gln	Phe	Ile	Gln	Arg	Ala	Ser	Thr	Ser	Gly	
166			225					230					235				
168	agg	ccc	ttc	ctg	ctc	tat	gtg	gct	ctg	gcc	cac	atg	cac	gtg	ccc	tta	1009
169	Arg	Pro	Phe	Leu	Leu	Tyr	Val	Ala	Leu	Ala	His	Met	His	Val	Pro	Leu	
170		240					245					250					
172	ccc	gtg	act	cag	cta	cca	gca	gcg	cca	cgg	ggc	aga	agc	ctg	tat	ggt	1057
173	Pro	Val	Thr	Gln	Leu	Pro	Ala	Ala	Pro	Arg	Gly	Arg	Ser	Leu	Tyr	Gly	
174	255					260					265				270		
176	gca	ggg	ctc	tgg	gag	atg	gac	agt	ctg	gtg	ggc	cag	atc	aag	gac	aaa	1105
177	Ala	Gly	Leu	Trp	Glu	Met	Asp	Ser	Leu	Val	Gly	Gln	Ile	Lys	Asp	Lys	

RAW SEQUENCE LISTING

DATE: 09/24/2001

PATENT APPLICATION: US/09/495,823

TIME: 12:05:32

Input Set : A:\5800-79.txt

Output Set: N:\CRF3\09242001\I495823.raw

178		275		280		285	
180	gtt gac cac aca	gtg aag gaa aac aca	ttc ctc tgg ttt aca	gga gac		1153	
181	Val Asp His Thr	Val Lys Glu Asn Thr	Phe Leu Trp Phe Thr	Gly Asp			
182		290		295		300	
184	aat ggc ccg tgg	gct cag aag tgt	gag cta gcg ggc	agt gtg ggt ccc		1201	
185	Asn Gly Pro Trp	Ala Gln Lys Cys Glu	Leu Ala Gly Ser	Val Gly Pro			
186		305		310		315	
188	ttc act gga ttt	tgg caa act cgt	caa ggg gga agt	cca gcc aag cag		1249	
189	Phe Thr Gly Phe	Trp Gln Thr Arg	Gln Gly Gly Ser	Pro Ala Lys Gln			
190		320		325		330	
192	acg acc tgg gaa	gga ggg cac cgg	gtc cca gca ctg	gct tac tgg cct		1297	
193	Thr Thr Trp Glu	Gly Gly His Arg	Val Pro Ala Leu	Ala Tyr Trp Pro			
194		335		340		345	350
196	ggc aga gtt cca	gtt aat gtc acc	agc act gcc ttg	tta agc gtg ctg		1345	
197	Gly Arg Val Pro	Val Asn Val Thr	Ser Thr Ala Leu	Leu Ser Val Leu			
198		355		360		365	
200	gac att ttt cca	act gtg gta gcc	ctg gcc cag gcc	agc tta cct caa		1393	
201	Asp Ile Phe Pro	Thr Val Val Ala	Leu Ala Gln Ala	Ser Leu Pro Gln			
202		370		375		380	
204	gga cgg cgc ttt	gat ggt gtg gac	gtc tcc gag gtg	ctc ttt ggc cgg		1441	
205	Gly Arg Arg Phe	Asp Gly Val Asp	Val Ser Glu Val	Leu Phe Gly Arg			
206		385		390		395	
208	tca cag cct ggg	cac agg gtg ctg	ttc cac ccc aac	agc ggg gca gct		1489	
209	Ser Gln Pro Gly	His Arg Val Leu	Phe His Pro Asn	Ser Gly Ala Ala			
210		400		405		410	
212	gga gag ttt gga	gcc ctg cag act	gtc cgc ctg gag	cgt tac aag gcc		1537	
213	Gly Glu Phe Gly	Ala Leu Gln Thr	Val Arg Leu Glu	Arg Tyr Lys Ala			
214		415		420		425	430
216	ttc tac att acc	ggt gga gcc agg	gcg tgt gat ggg	agc acg ggg cct		1585	
217	Phe Tyr Ile Thr	Gly Gly Ala Arg	Ala Cys Asp Gly	Ser Thr Gly Pro			
218		435		440		445	
220	gag ctg cag cat	aag ttt cct ctg	att ttc aac ctg	gaa gac gat acc		1633	
221	Glu Leu Gln His	Lys Phe Pro Leu	Ile Phe Asn Leu	Glu Asp Asp Thr			
222		450		455		460	
224	gca gaa gct gtg	ccc cta gaa aga	ggt ggt gcg gag	tac cag gct gtg		1681	
225	Ala Glu Ala Val	Pro Leu Glu Arg	Gly Gly Ala Glu	Tyr Gln Ala Val			
226		465		470		475	
228	ctg ccc gag gtc	aga aag gtt ctt	gca gac gtc ctc	caa gac att gcc		1729	
229	Leu Pro Glu Val	Arg Lys Val Leu	Ala Asp Val Leu	Gln Asp Ile Ala			
230		480		485		490	
232	aac gac aac atc	tcc agc gca gat	tac act cag gac	cct tca gta act		1777	
233	Asn Asp Asn Ile	Ser Ser Ala Asp	Tyr Thr Gln Asp	Pro Ser Val Thr			
234		495		500		505	510
236	ccc tgc tgt aat	ccc tac caa att	gcc tgc cgc tgt	caa gcc gca taa		1825	
237	Pro Cys Cys Asn	Pro Tyr Gln Ile	Ala Cys Arg Cys	Gln Ala Ala *			
238		515		520		525	
240	cagaccaatt tttattccac	gaggaggagt acctggaaat	taggcaagtt tgcttccaaa			1885	
241	tttcattttt accctcttta	caaacacacg cttagtttta	gtcttgaggat ttagtttttg			1945	
242	agttagcctt gcatatccct	tctgtatcct gtccctcctc	cacgccgacc cgagagcagc			2005	

RAW SEQUENCE LISTING

DATE: 09/24/2001

PATENT APPLICATION: US/09/495,823

TIME: 12:05:32

Input Set : A:\5800-79.txt

Output Set: N:\CRF3\09242001\I495823.raw

```

243 tgagctgcgc tggctctggg cagggagtgt gccttaatgg gaagcacacg ggctttggag 2065
244 tcaggcacag gtgccagctc cagcttttga acttgggcaa ttgtttaacc taacctgcaa 2125
245 gttgattttg agggttaaat aaaggcatac atgaaaaaaaa aaaaaaaaaa 2175
247 <210> SEQ ID NO: 3
248 <211> LENGTH: 871
249 <212> TYPE: PRT
250 <213> ORGANISM: homo sapiens
252 <400> SEQUENCE: 3
253 Met Lys Tyr Ser Cys Cys Ala Leu Val Leu Ala Val Leu Gly Thr Glu
254 1 5 10 15
255 Leu Leu Gly Ser Leu Cys Ser Thr Val Arg Ser Pro Arg Phe Arg Gly
256 20 25 30
257 Arg Ile Gln Gln Glu Arg Lys Asn Ile Arg Pro Asn Ile Ile Leu Val
258 35 40 45
259 Leu Thr Asp Asp Gln Asp Val Glu Leu Gly Ser Leu Gln Val Met Asn
260 50 55 60
261 Lys Thr Arg Lys Ile Met Glu His Gly Gly Ala Thr Phe Ile Asn Ala
262 65 70 75 80
263 Phe Val Thr Thr Pro Met Cys Cys Pro Ser Arg Ser Ser Met Leu Thr
264 85 90 95
265 Gly Lys Tyr Val His Asn His Asn Val Tyr Thr Asn Asn Glu Asn Cys
266 100 105 110
267 Ser Ser Pro Ser Trp Gln Ala Met His Glu Pro Arg Thr Phe Ala Val
268 115 120 125
269 Tyr Leu Asn Asn Thr Gly Tyr Arg Thr Ala Phe Phe Gly Lys Tyr Leu
270 130 135 140
271 Asn Glu Tyr Asn Gly Ser Tyr Ile Pro Pro Gly Trp Arg Glu Trp Leu
272 145 150 155 160
273 Gly Leu Ile Lys Asn Ser Arg Phe Tyr Asn Tyr Thr Val Cys Arg Asn
274 165 170 175
275 Gly Ile Lys Glu Lys His Gly Phe Asp Tyr Ala Lys Asp Tyr Phe Thr
276 180 185 190
277 Asp Leu Ile Thr Asn Glu Ser Ile Asn Tyr Phe Lys Met Ser Lys Arg
278 195 200 205
279 Met Tyr Pro His Arg Pro Val Met Met Val Ile Ser His Ala Ala Pro
280 210 215 220
281 His Gly Pro Glu Asp Ser Ala Pro Gln Phe Ser Lys Leu Tyr Pro Asn
282 225 230 235 240
283 Ala Ser Gln His Ile Thr Pro Ser Tyr Asn Tyr Ala Pro Asn Met Asp
284 245 250 255
285 Lys His Trp Ile Met Gln Tyr Thr Gly Pro Met Leu Pro Ile His Met
286 260 265 270
287 Glu Phe Thr Asn Ile Leu Gln Arg Lys Arg Leu Gln Thr Leu Met Ser
288 275 280 285
289 Val Asp Asp Ser Val Glu Arg Leu Tyr Asn Met Leu Val Glu Thr Gly
290 290 295 300
291 Glu Leu Glu Asn Thr Tyr Ile Ile Tyr Thr Ala Asp His Gly Tyr His
292 305 310 315 320
293 Ile Gly Gln Phe Gly Leu Val Lys Gly Lys Ser Met Pro Tyr Asp Phe

```

Use of n and / or Xaa has been detected in the
Sequence Listing. Review the Sequence Listing
to ensure a corresponding explanation is present
in the <220> to <223> fields of each sequence
using n or Xaa.

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/495,823

DATE: 09/24/2001

TIME: 12:05:33

Input Set : A:\5800-79.txt

Output Set: N:\CRF3\09242001\I495823.raw

L:626 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4